USING ARTIFICIAL INTELLIGENCE APPROACH IN PREDICTING CHLORIDE CONTENT IN CONCRETE EXPOSED TO CHLORIDE SOLUTION

Nguyen Huu Nam[1], Tran Quoc Thanh[1]
Giap Van Loi[2], Tran Kim Phuong[2], Tran Van Quan[2]
1. Hydropower Institute & Renewables, Vietnam Academy for Water Resources
2. University of Transport and Technology

Abstract: *Structural corrosion due to chloride ingress is considered to be the greatest threat to buildings. Structures such as bridges, roads, infrastructure, and harbors are exposed to chloride-rich environments, either through anti-freeze salting or from the natural environment. In the marine environment, tidal and wave areas are considered to be at high risk of corrosion. The corrosion process is highly dependent on the chloride concentration in the concrete structure. Therefore, the determination of chloride concentration in concrete is meaningful and reliable. In this study, a Gradient Boosting (GB) machine learning model is proposed to predict chloride content in concrete. A dataset of 325 experimental results was collected from international literature. The results of the model are programmed and run in the Python platform, the authors use the tool of the Gradient Boosting model to predict the output parameters of chloride content in concrete. The results show that the proposed model can predict the chloride content in concrete simply and quickly, helping design engineers to predict the chloride content in advance based on the input parameters.*

Keywords: *Artificial Intelligence (AI), Gradient Boosting (RF), chloride content, concrete, Reinforcement corrosion.*

1. INTRODUCTION

Chloride may be present in concrete constituents, especially sand and water. According to the Standard of European EN 206-1, chloride ion content to a mass of cement shall not exceed 0.4% for reinforced concrete or concrete with steel core, 0.2% for prestressed concrete, and 1% for unreinforced concrete. This standard also does not allow the use of chloride-based additives when using reinforced concrete or steel core concrete.

The concentration and contact environment of chloride greatly affect the corrosion process as well as the service life of concrete, in addition, it also depends on other factors such as the ratio of water/cement, the content of aggregates. [2], [3], [4]. Until now, much research works have been done to predict the chloride content in concrete and investigate the influence of chloride under different conditions on the corrosion process in concrete. [5], [6]. Most of these studies have given predictive equations based on the corrosion mechanisms of concrete [7], [8]. These methods have introduced the corrosion mechanism and have not mentioned the corrosion process of concrete under different working environmental conditions such as the submerged condition of the structure. There are no studies to predict chloride concentration in concrete structures under corrosive conditions.

With the development of computer science, artificial intelligence will solve engineering models to determine the life of these parameter

Receipt Date: September 12th, 2022
Review Approval Date: October 28th, 2022
Publish Approval Date: November 11th, 2022

structures on [4]-[9], these models will solve complex problems based on data about concrete properties and working environment. In the past time, many mathematical models have been developed to predict the chloride content in concrete such as Gradient Boosting network model, artificial neural network (ANN),... In the recent period, the prominent method for prediction and widely used in the construction industry is the Gradient Boosting network model. This research objective is based on a data set of 325 experimental results published in a prestigious scientific journal with the input parameters being contact environment, ratio water/cement, silica fume content, and the output is the chloride content in the concrete. It provides construction engineers with documents that allow for the evaluation of the chloride content in concrete, serving as a basis for quickly and accurately assessing the corrosion process in concrete.

2. MODEL SETTING

2.1. Gradient Boosting algorithm

Gradient Boosting algorithm was first introduced in 1990 by machine learning. By iterating over a few simple patterns, instead of trying to build the best model, we build a family of weaker models that, when combined, produce a superior model [10]. In the first stage, Boost was designed as a machine learning process with the aim of improving binary outcome prediction, implemented through the learner base (using decision tree weights) [11]. The enhancement is then considered to reduce gradients in the function space and can be used to fit regression models.

Gradient Boosting (GB) has advantages over other statistical learning methods in that it can interpret results, simplifies complex interactions, and also requires data pre-processing and parameter adjustment [12]. In addition, GB increases speed and improves model and selection accuracy. In this study, Gradient boost used a fixed-size decision tree model as the basis and used boost algorithms developed by Jerome H. Friedman [13], [14].

2.2. Evaluate model predictability

To evaluate the accuracy of the built GB model, three criteria are used in this study, the correlation coefficient (R) and the Root Mean Square Error (RMSE). The R-index represents the correlation between the predicted outcome and the actual output, with a value ranging from 0 to 1. The higher the value of R, the better the correlation between the predicted value and the value reality. The RMSE and MAE indices are another method of error determination, based on the mean squared difference between the predicted output and the actual output. Low RMSE and MAE values indicate better performance of the ANN algorithm. Formulas for calculating R, RMSE and MAE can be found in the citation [15].

2.3. Data used

In this study, 325 actual data on chloride concentration were collected from documents published in prestigious international journals [6]. Gradient Boosting model uses 4 input variables including: (1) submerged depth, (2) water/cement ratio, (3) silica fume content. The output parameter to be considered is the chloride content of the sample. The dataset used in this work was randomly split into two sub-datasets using a uniform distribution, where 70% of the data was used to train the GB models and 30% of the data is used for model validation. The description of the data is shown in Figure 1. The test samples were subjected to different submerged environmental conditions in Fig 1.a including: (1) fully submerged environment, (2) waved concrete sample environment, (3) environment above sea water is not flooded and waves. Fig 1.a shows the chloride content in the samples evenly distributed in the three types of samples above with a depth of 5 to 30mm.

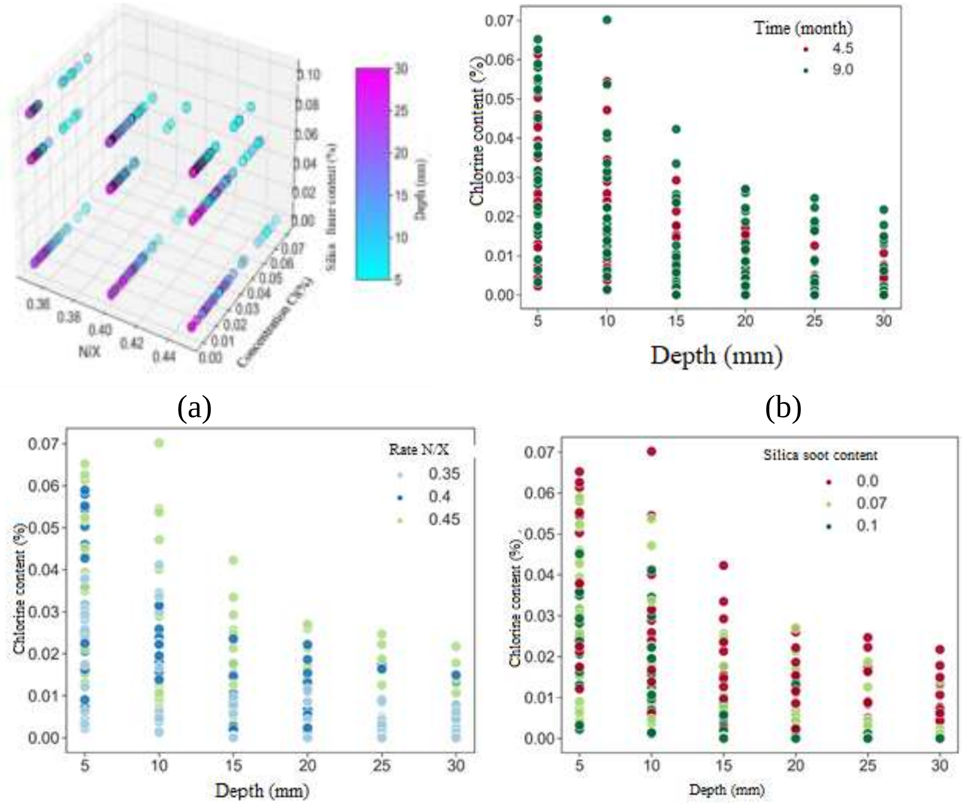


Fig 1: Relationship between input and output variables

3. RESULTS AND DISCUSSION

The optimal GB model defined in this section allows us to predict the chloride content in the concrete sample for training and testing. Fig 2 and Fig 3 are the results of the predicted chloride content for the training part and the corresponding error frequency between the model's predicted value and the actual value, respectively. Fig 4 and Fig 5 are the results of the predicted chloride content for the control and the corresponding error frequency between the predicted value of the GB model for the control and the actual value, respectively.

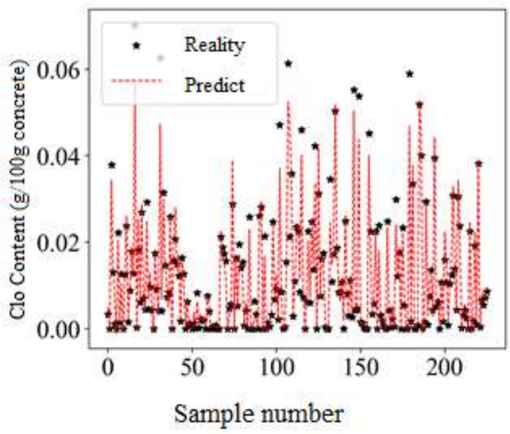


Fig 2: Chloride content predicted for the training part by the model GB

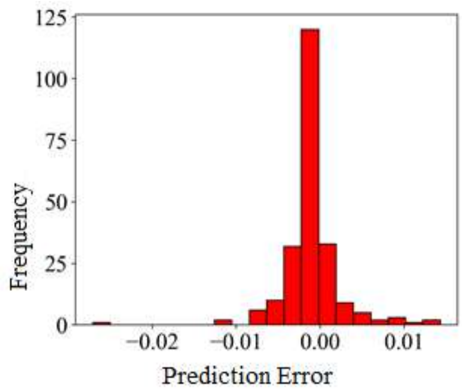


Fig 3: Frequency of error between the predicted chloride content by GB and the actual value for the training

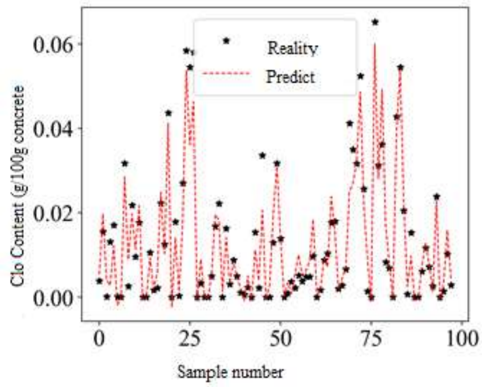


Fig 4: Chloride content predicted for the testing by model GB

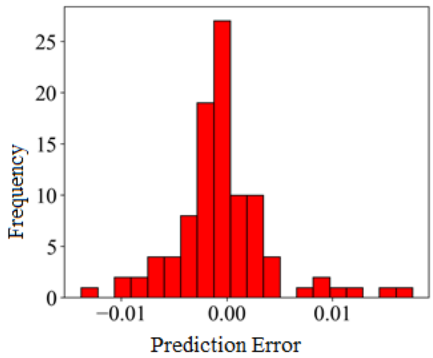
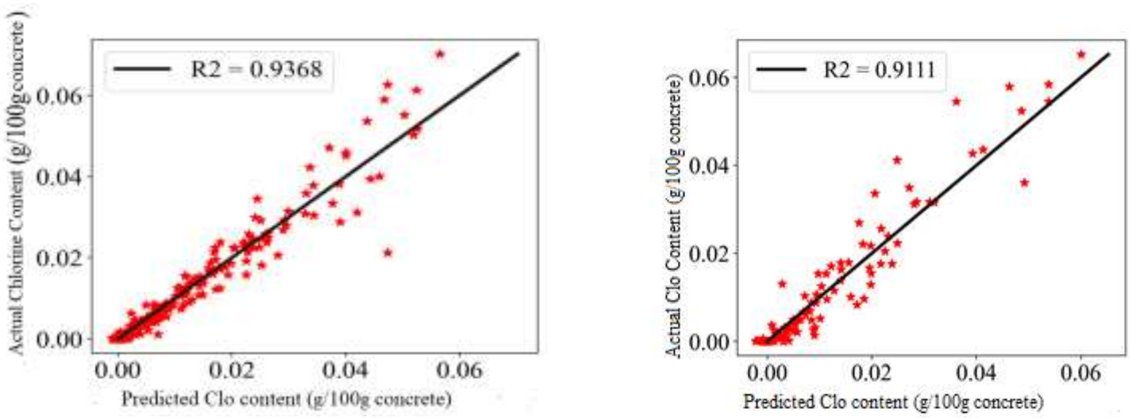


Fig 5: Frequency of error between the chloride content predicted by GB and the actual value for the training

The results show that the Gradient boosting model predicts the chloride content for 227 concrete samples with relatively high accuracy for the training and testing part. Figures 3 and 5 show the frequency of error between the predicted chloride content and the actual value for the training part with the range [-0.03; 0.015] (g/100g concrete) is higher than the control in the range [-0.015; 0.02] (g/100g concrete). However, the prediction errors of the prediction and testing part with a large number of samples are concentrated at the value 0. These errors show that the prediction ability of the proposed Gradient Boosting model is feasible with low errors.

The regression model for 2 parts of training and testing is shown in Figure 6. From that, we can see that the predictability of the model is relatively high, quite close to the actual friction angle, but there is still sample error with high chloride content. The coefficient of determination R^2 obtained for the training part is R^2=0.9368 and the test part is R^2=0.9111. It shows that the value between the training and the testing is almost equivalent, so it is very possible to apply the Gradient booting model to predict the chloride content for the concrete sample.

Especially with the Gradient Booting model, the important role of each component in the chloride content is also shown in Figure 8. The results in Figure 8 shows that the factor that has the greatest influence on the chloride content is the depth of the sample, followed by the medium and the silica fume content and the ratio of water/cement and only time are the factors that have the smallest influence on the chloride content in concrete.



(a) (b)

Fig 6: Result of model regression gradient boosting (a) Training, (b) Testing

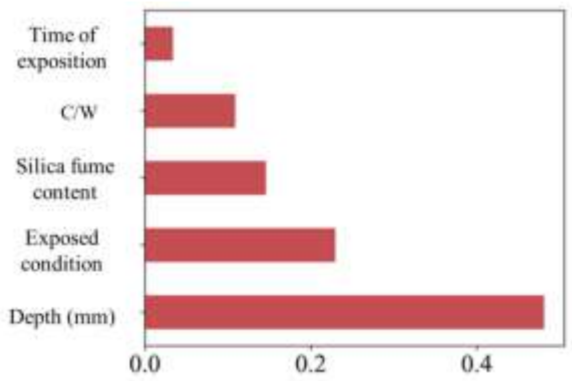


Fig 7: Factors affecting chloride content in concrete analyzed by gradient boosting model

4. CONCLUSIONS

In this study, a machine learning model of artificial intelligence (AI) in chloride content prediction was tested. To build and develop the model, a data set of 325 experimental results has been collected from prestigious journals in the world, including 5 input data parameters and 1 output parameter of chloride content in concrete. Test concrete samples are a collection of many different concrete samples working in different flooded and environmental conditions, so there is variation. Therefore, the application of the Gradient Boosting model, an artificial intelligence algorithm, to predict and analyze the chloride content in concrete holds both scientific and practical importance. It enhances the predictability of reinforcement corrosion in concrete and the structural bearing capacity. The results show that the Gradient Boosting model is feasible in determining the chloride content in the remaining 100g of concrete of the deformed soil with the coefficient of determination for the training model of 0.9368 and the control of 0.9111. The error of the training model applied to the test is very small. The results show that the application of the proposed Gradient Boosting model is capable of accurately predicting the chloride content in concrete, minimizing the cost and time.

REFERENCES

[1] K. Tuutti, ‘Corrosion of steel in concrete’, Swedish Cement and Concrete Research Institute, Stockholm, 1982.

[2] ‘Chloride Profiles – Analysis and Interpretation.pdf’.

[3] A. Khitab, W. Anwar, and M. T. Arshad, ‘Predictive Models of Chloride Penetration in concrete: An Overview’, *MUST J. Eng. and Appl. Sci.*, vol. 1, no. 1, pp. 1–14, Mar. 2017, doi: 10.22496/mjeas20170103.

[4] N. D. Stambaugh, T. L. Bergman, and W. V. Srubar, ‘Numerical service-life modeling of chloride-induced corrosion in recycled-aggregate concrete’, *Construction and Building Materials*, vol. 161, pp. 236–245, Feb. 2018, doi: 10.1016/j.conbuildmat.2017.11.084.

[5] V. Chandwani, V. Agrawal, and R. Nagar, ‘Modeling slump of ready mix concrete using genetic algorithms assisted training of Artificial Neural Networks’, *Expert Systems with Applications*, vol. 42, no. 2, pp. 885–893, Feb. 2015, doi: 10.1016/j.eswa.2014.08.048.

[6] N. Ukrainczyk and V. Ukrainczyk, ‘A neural network method for analysing concrete durability’, *Magazine of Concrete Research*, vol. 60, no. 7, pp. 475–486, Sep. 2008, doi: 10.1680/macr.2007.00016.

[7] E. Güneyisi, M. Gesoğlu, T. Özturan, and E. Özbay, ‘Estimation of chloride permeability of concretes by empirical modeling: Considering effects of cement type, curing condition and age’, *Construction and Building Materials*, vol. 23, no. 1, pp. 469–481, Jan. 2009, doi: 10.1016/j.conbuildmat.2007.10.022.

[8] L. Kong and X. Chen, ‘Influence mechanism of lightweight aggregate on concrete impermeability: prediction by ANN’, *Magazine of Concrete Research*, vol. 67, no. 1, pp. 17–26, Jan. 2015, doi: 10.1680/macr.14.00127.

[9] W. V. Srubar, ‘Stochastic service-life modeling of chloride-induced corrosion in recycled-aggregate concrete’, *Cement and Concrete Composites*, vol. 55, pp. 103–111, Jan. 2015, doi: 10.1016/j.cemconcomp.2014.09.003.

[10] S. Touzani, J. Granderson, and S. Fernandes, ‘Gradient boosting machine for modeling the energy consumption of commercial buildings’, *Energy and Buildings*, vol. 158, pp. 1533–1543, Jan. 2018, doi: 10.1016/j.enbuild.2017.11.039.

[11] R. E. Schapire, ‘Using output codes to boost multiclass learning problems’, p. 9.

[12] L. Guelman, ‘Gradient boosting trees for auto insurance loss cost modeling and prediction’, *Expert Systems with Applications*, vol. 39, no. 3, pp. 3659–3667, Feb. 2012, doi: 10.1016/j.eswa.2011.09.058.

[13] J. H. Friedman, ‘Greedy Function Approximation: A Gradient Boosting Machine’, *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.

[14] J. H. Friedman, ‘Stochastic gradient boosting’, *Computational Statistics & Data Analysis*, vol. 38, no. 4, pp. 367–378, Feb. 2002, doi: 10.1016/S0167-9473(01)00065-2.

[15] H. Nguyen, H.-B. Ly, T. Van Quan, T.-A. Nguyen, T.-T. Le, and B. Pham, ‘Optimization of Artificial Intelligence System by Evolutionary Algorithm for Prediction of Axial Capacity of Rectangular Concrete Filled Steel Tubes under Compression’, *Materials*, vol. 13, Mar. 2020, doi: 10.3390/ma13051205.

[16] S. K. Das, P. Samui, S. Z. Khan, and N. Sivakugan, ‘Machine learning techniques applied to prediction of residual strength of clay’, *cent.eur.j.geo.*, vol. 3, no. 4, pp. 449–461, Dec. 2011, doi: 10.2478/s13533-011-0043-1.

[17] S. Z. Khan, S. Suman, M. Pavani, and S. K. Das, ‘Prediction of the residual strength of clay using functional networks’, *Geoscience Frontiers*, vol. 7, no. 1, pp. 67–74, Jan. 2016, doi: 10.1016/j.gsf.2014.12.008.